Beginning of 08/27/2025

Contents today: (i) pseudorandom property — expander mixing lemma, (ii) probabilistic property — rapid mixing, and (iii): applications — Cheeger's inequality and sparsest cut.

Pseudorandom property

We want to compare d-regular expanders with random d-regular graphs (uniformly sample d edges incident on each vertex). The quantity we are interested in is $E(S,T)$, the set of edges between $S, T \subset G$.
In a random d-regular random graph G, $\mathbb{E}|E(S,T)| = |S| \cdot |T| \cdot d/n$. We include multiplicity — if $S \cap T$ overlap, an edge with both endpoints in the overlap will be counted twice.
If we instead look at d-regular expander, the pseudorandomness is in the following sense.

Proposition: Expander Mixing Lemma

If G is a two-sided d-regular expander, then for all $S, T \subset V$,

$$\left| |E(S,T)| - \frac{d}{n}|S| \cdot |T| \right| \leqslant \lambda d \sqrt{|S| \cdot |T|}$$

Proof. By d-regularity, the normalized $\overline{A}$ is just A/d. Then

$$|E(S,T)| = d \cdot \langle \mathbf{1}_S, \overline{A}\mathbf{1}_T \rangle = d \cdot \left\langle \mathbf{1}_S, \sum_{i=1}^{n} \lambda_i v_i v_i^T \mathbf{1}_T \right\rangle.$$

The first eigenvalue is 1 with eigenvector $\mathbf{1}/\sqrt{n}$ so that the first term $d \langle \mathbf{1}_S, \lambda_1 v_1 v_1^T \mathbf{1}_T \rangle = d \cdot |S| \cdot |T|/n$.
Thus

$$\begin{aligned} \left| |E(S,T)| - \frac{d|S||T|}{n} \right| &= \left| d \left\langle \mathbf{1}_S, \sum_{i \geqslant 2} \lambda_i v_i v_i^T \mathbf{1}_T \right\rangle \right| \\ &\leqslant d\lambda \|\mathbf{1}_S\|_2 \|\mathbf{1}_T\|_2 = \lambda d \sqrt{|S| \cdot |T|}. \end{aligned}$$

□

Remark. A similar claim holds even if we drop the d-regularity assumption (only need two-sided λ-expander): $|E(S,T)| = \langle D^{-1/2}\mathbf{1}_S, \overline{A} D^{1/2} \mathbf{1}_T \rangle$. The contribution of v_1 (recall $v_1 = \mathbf{1}D^{1/2}$) is

$$\langle D^{-1/2}\mathbf{1}_S, v_1 v_1^T D^{1/2} \mathbf{1}_T \rangle = \langle D^{1/2}\mathbf{1}_S, D^{1/2}\mathbf{1}\mathbf{1}D^{1/2}D^{1/2}\mathbf{1}_T \rangle = \frac{\mathrm{vol}(S)\mathrm{vol}(T)}{\mathrm{vol}(V)}.$$

We skip the rest and we claim that the comparison is made between a random 2-sided λ-expander against a random graph with identical degree profile. Becuase the graph is no longer regular, we replace $\lambda d \sqrt{|S| \cdot |T|}$ by $\lambda \sqrt{\mathrm{vol}(S) \cdot \mathrm{vol}(T)}$.
Another remark. Given a degree profile, how do we sample graphs following this? We start by drawing a bunch of "half edges" around each vertex, and then do matching. Under this perspective, $\mathrm{vol}(S)$ is the number of half-edges around S, and $\mathrm{vol}(S)\mathrm{vol}(T)$ is the total number of "valid" matchings between them, which is then divided by $\mathrm{vol}(V)$, the "total" number of "valid" matchings.
Final remark. Two-sided expanders are necessary. Consider a bipartite graph $G = (U, V, E)$ where U, V are the partitions. It is known that the spectrum of a bipartite matrix is symmetric around 0. Hence $\lambda_1 = 1$ implies $\lambda_n = -1$ so it cannot be two-sided λ-expanders.

If S,T are chosen from the same side, $E(S,T) = \varnothing$ and counterexamples can arise. There does exist a variant **expander mixing lemma for bipartite graphs**: for all $S \subset U, T \subset V$,

$$\left| |E(S,T)| - \frac{\mathrm{vol}(S)\cdot\mathrm{vol}(T)}{\mathrm{vol}(V)} \right| \leqslant \lambda\sqrt{\mathrm{vol}(S)\cdot\mathrm{vol}(T)}.$$

Probabilistic property

We define a random walk operator P on G with $P = AD^{-1}$ (column stochastic). We also let π denote the stationary distribution, i.e., $\pi P = \pi$. This requires $\pi(v) = d(v)/\mathrm{vol}(V)$.
Given two distributions μ, π, the **total variation distance** (TVD) between them is defined by $\|\mu-\pi\|_{\mathrm{TV}} = \|\mu-\pi\|_1/2$.
We say that a random walk P **mixes** if for all distribution over V,

$$\lim_{t\to\infty} \|P^t\mu - v\|_{\mathrm{TV}} = 0.$$

We say P mixes **rapidly** if the decay above is exponential (since it's a property of P, this bound needs to be uniform over all μ). When the context is clear, we drop the subscript $\|\cdot\|_{\mathrm{TV}}$.

Proposition

Let G be a two-sided λ-expander. Then for all μ,

$$\|P^t\mu - \pi\| \leqslant \frac{1}{2}\sqrt{1/\pi_{\min}}\cdot\lambda^t$$

where $\pi_{\min} = \min_{a\in V}\pi(a)$. This bound is less tight for small t and becomes tighter over larger t's.

Proof. Since $P = AD^{-1}$, we have $P = D^{1/2}\overline{A}D^{-1/2}$ so P and $\overline{A}$ are **similar** and in particular have the identical spectrums, with eigenvectors u_i of D as $D^{1/2}\cdot v_i$ (eigenvectors of A).
A small technical reduction: since $\|\|\cdot_1$ is convex, so is TVD, and thus it suffices to consider $\mu = \mathbf{1}_v$ (one-hot distribution) for all $v \in V$.
Now fix some v and consider $P^t\mathbf{1}_v$. Observe

$$P^t\mathbf{1}_v = D^{1/2}[\overline{A}]^t D^{-1/2}\mathbf{1}_v.$$

For now we focus on $[\overline{A}]^t D^{-1/2}\mathbf{1}_v$. Noting that $v_1 = D^{1/2}\mathbf{1}/\sqrt{\mathrm{vol}(V)}$, we have by eigenvalue expansion

$$\begin{aligned}[\overline{A}]^t D^{-1/2}\mathbf{1}_v &= [\overline{A}]^t\left(\frac{\langle D^{1/2}\mathbf{1}, D^{-1/2}\mathbf{1}_v\rangle}{\sqrt{\mathrm{vol}(V)}} + \text{something}\right)\\ &= \frac{1}{\mathrm{vol}(V)}\cdot\mathrm{diag}(1/\sqrt{d(v)}) + [\overline{A}]\cdot\text{something}.\end{aligned}$$

On the other hand, using $D^{-1/2}\pi = \sqrt{1/\mathrm{vol}(V)}v_1$, most of the terms cancel out:

$$P^t\mathbf{1}_v - \pi = D^{1/2}([\overline{A}]^t D^{-1/2}\mathbf{1}_v - D^{-1/2}\pi) = D^{1/2}[\overline{A}]^t\cdot\text{something}$$

where "something" lives in the space orthogonal to v_1. Call it $\vec{\epsilon}$. To be continued next time: $\|P^t\mathbf{1}_v - \pi\|_1/2 = \|D^{1/2}[\overline{A}]^t\vec{\epsilon}\|_1/2$ and use C-S to obtain the original term $1/2\cdot\sqrt{1/\pi_{\min}}\cdot\lambda^t$. ☐